

# Discussion Papers

# 991

Joachim R. Frick • Markus M. Grabka • Olaf Groh-Samberg

## Dealing with Incomplete Household Panel Data in Inequality Research

Berlin, March 2010

Opinions expressed in this paper are those of the author(s) and do not necessarily reflect views of the institute.

## IMPRESSUM

© DIW Berlin, 2010

DIW Berlin  
German Institute for Economic Research  
Mohrenstr. 58  
10117 Berlin  
Tel. +49 (30) 897 89-0  
Fax +49 (30) 897 89-200  
<http://www.diw.de>

ISSN print edition 1433-0210  
ISSN electronic edition 1619-4535

Available for free downloading from the DIW Berlin website.

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN.  
Papers can be downloaded free of charge from the following websites:

[http://www.diw.de/de/diw\\_01.c.100406.de/publikationen\\_veranstaltungen/publikationen/diskussionspapiere/diskussionspapiere.html](http://www.diw.de/de/diw_01.c.100406.de/publikationen_veranstaltungen/publikationen/diskussionspapiere/diskussionspapiere.html)  
<http://ideas.repec.org/s/diw/diwwpp.html>  
[http://papers.ssrn.com/sol3/JELJOUR\\_Results.cfm?form\\_name=journalbrowse&journal\\_id=1079991](http://papers.ssrn.com/sol3/JELJOUR_Results.cfm?form_name=journalbrowse&journal_id=1079991)

# **Dealing with incomplete household panel data in inequality research**

**Joachim R. Frick\***  
**Markus M. Grabka**  
**Olaf Groh-Samberg**

Berlin, March 2010

*\* Corresponding author:*

PD Dr. Joachim R. Frick, [jfrick@diw.de](mailto:jfrick@diw.de)  
SOEP at DIW Berlin, TU Berlin, IZA Bonn  
Mohrenstrasse 58, 10117 Berlin, Germany  
phone: +49-30-89789-279, Fax: +49-30-89789-109

Dr. Markus M. Grabka, [<mgrabka@diw.de>](mailto:mgrabka@diw.de)  
SOEP at DIW Berlin; TU Berlin

Prof. Dr. Olaf Groh-Samberg, [ogrohsamberg@bigsss.uni-bremen.de](mailto:ogrohsamberg@bigsss.uni-bremen.de)  
BIGSSS, University of Bremen, DIW Berlin

## Abstract:

Population surveys around the world face the problem of declining cooperation and participation rates of respondents. Not only can item nonresponse and unit nonresponse impair important outcome measures for inequality research such as total household disposable income; there is also a further case of missingness confronting household panel surveys that potentially biases results. The approach commonly used in such surveys of interviewing *all* adult household members and aggregating their individual incomes to yield a final outcome measure for welfare analyses often suffers from partial unit non-response (PUNR), i.e., the non-response of at least one unit, or member, of an otherwise participating household. In these cases, the *aggregate income* of all household members lacks at least one individual's income. These processes are typically not random and require appropriate correction.

Using data from more than twenty waves of the German Socio-Economic Panel (SOEP) we evaluate four different strategies to deal with this phenomenon: (a) *Ignorance*, i.e., assuming the missing individual's income to be zero. (b) *Adjustment* of the equivalence scale to account for differences in household size and composition. (c) *Elimination* of all households observed to suffer PUNR, and re-weighting of households observed to be at risk of but not affected by PUNR. (d) *Longitudinal imputation* of the missing income components. The aim of this paper is to show how the choice of technique affects substantive results in the inequality research. We find indications of substantial bias on income inequality and poverty as well as on income mobility. These findings are obviously even more important in cross-national comparative analyses if the data providers in the individual countries deal differently with PUNR in the underlying data.

*Keywords:* Household Panel Surveys, Partial Unit Non-Response, Inequality, Mobility, Imputation, SOEP

*JEL-Codes:* D33, I32, C81

# 1 Motivation

One of the standard assumptions of welfare economics is that individuals living together in a “needs unit”—usually a private household—pool and share all their available resources. This approach, which is generally accepted in the inequality research, is not just a means of adapting to data limitations: it is based on the idea that household members seek to achieve a common and equal standard of living (e.g., Canberra Group 2001, Atkinson & Bourgignon 2000, Smeeding & Weinberg 2001).<sup>i</sup> This requires that all incomes received in a given household are aggregated across all members and that the total sum is distributed among all of them. Typically, an equivalence scale is then applied in order to adjust for differences in household composition and size, thus allowing for economies of scale in larger households as well as variation in needs across age groups (see, e.g., Buhmann et al 1988). This approach crucially depends on either one household representative (e.g., the household head) providing complete proxy information on behalf of all members (similar to the approach used in the US Panel Study of Income Dynamics, PSID, which was started in 1968) or, as is common in almost all household panel surveys started thereafter, all adult household members actually providing an interview themselves. Various sorts of non-response behavior appear in most of these surveys, however, posing a serious threat to the implicit assumption of (representative) full coverage of all the resources and needs of individuals living together in the household<sup>ii</sup>. In such cases, unit non-response (UNR, i.e., the non-response of an entire household) is addressed by means of proportionally weighting successfully surveyed observations, while item non-response (INR) is corrected for by either weighting or imputation. But there remains the problem of partial unit-non response (PUNR) in those households where at least one member does not participate while other members do. In fact, there appears to be a generally increasing percentage of households affected by partial unit non-response (PUNR) in population surveys. Ignorance of this phenomenon may give rise to several problems: a) misreporting income aggregates; b) increasing bias in results on income inequality, poverty, and mobility, as well as c) bias in analyses on the intra-household distribution.

In the literature on social welfare, various approaches have been proposed for dealing with this phenomenon in the generation of “equivalent household income,” the outcome variable relevant for inequality and poverty analysis: (a) Ignoring the fact that a household member (and his/her income information) is missing, thus assuming the non-responding individual’s income is zero. This does not rule out that the person is living entirely from household-

based transfers (public and private), which can be assumed to be measured correctly and comprehensively based on the information provided by other household members. (b) Adjusting the calculation of the equivalence scale to ignore the person's share in household needs in order to compensate for their missing income. This approach implicitly assumes that the incomes of other household members are independent of the income of the missing individual. (c) Eliminating all households observed with PUNR from the analysis population, thus assuming that there are no systematic differences between them and completely surveyed households, i.e., that the underlying missing process is completely at random. An extension of this approach would try to compensate for potential selectivity by means of weighting. This can be accomplished by proportionally increasing the share of those at risk of but not affected by PUNR. That is, for all households which are not at risk of PUNR because there is only one adult respondent, the weighting factor remains unchanged. Obviously, this weighting strategy can be more or less complex depending on the case at hand. (d) Imputing the missing income components at the individual level and aggregating across all household members (including those with PUNR).

At first glance, the assumptions of options (a) through (c) are very strong given the (likely) selectivity issues involved in the missing mechanism. Options (c) [with re-weighting] and (d) appear to be less selective, and, in principle, option (d) has the additional advantage of maintaining the entire survey population. Having said that, there is quite some normativity involved in the actual implementation of any such imputation process, as is true in any case of imputation. For example, one may argue that the degree of misspecification is actually a general underreporting that can be corrected for by either adjusting the household income by means of a “(relative) factor” or by adding an “(absolute) flat sum.” More appropriately, one may allow for more variation with respect to the contribution of various income components to the overall household income measure and by controlling for household and individual characteristics related to the missing mechanism. For example, the severity of misreporting is probably very different for a household where data on the 85-year-old mother of the household head is missing due to lingering illness than for a household where the main income-earner is absent because he/she is drilling for oil on an offshore platform.

The results from research on equivalent household income inequality and especially on relative poverty may crucially depend on the choice of the aforementioned option, because this decision will affect (a) the income of all individuals living in households affected by PUNR, and (b) the relative poverty line to be derived from the national mean or median in-

come. Having said that, any bias in cross-sectional income measures will (c) also affect mobility analyses based on those measures.

Using more than twenty waves of microdata from the German Socio-Economic Panel (SOEP), this paper assesses how approaches to PUNR can potentially impact income inequality, income poverty, and mobility. The paper is set up as follows: In Section 2, we first describe the incidence of and trends in PUNR in the SOEP over the period 1984 to 2007, before turning to the analysis of selectivity of PUNR. Here we control for the relevance of concurrent household characteristics (e.g., size and composition) and for individual characteristics of the missing person. Section 3 presents the various techniques for dealing with PUNR, focusing on the principles of our three-stage imputation strategy for income components at the individual level. In Section 4, we provide sensitivity analyses showing the variation in the results for income inequality and poverty when choosing among the aforementioned options. Making explicit use of the panel nature of the underlying data, we also demonstrate the importance of PUNR (and how it is treated in the microdata) for poverty dynamics and income mobility. Here, our findings show that dynamics are exaggerated if PUNR is present in at least one wave. Section 5 concludes with some remarks on the potential relevance of our findings for cross-national comparability of research on income inequality, poverty, and mobility.

## **2 Incidence and selectivity of PUNR in the German Socio-Economic Panel (SOEP)**

### **2.1 The data**

The German Socio-Economic Panel (SOEP) is a representative longitudinal survey of individuals living in private households in Germany (Wagner et al. 2007). The survey was started in 1984 in West Germany and was extended to East Germany in June 1990, somewhat more than half a year after the fall of the Berlin Wall. The initial sample included over 12,000 respondents, with everyone aged 17 and over in sample households being interviewed. In recent years, several representative new sub-samples have been drawn, which have approximately doubled the initial sample size. Other additional samples were explicitly designed to sample specific subgroups of the population. In 1995, the SOEP introduced an oversampling of immigrants to cope with the misrepresentation of recent immigrants in ongoing panel surveys. In 2002, to overcome the problem of a lack of information on “rich” individuals in rep-

representative population surveys—an important group for welfare analyses given the high concentration of economic resources (income and wealth) at the top of the distribution—the SOEP introduced a high-income sample overrepresenting the top 3% of the income distribution. The sample analyzed below employs all available observation years up to survey year 2007.

One of the main problems population surveys face when asking for (specific) income and wealth information is non-response, and SOEP is no exception. To make effective use of the panel nature of SOEP, all cases of item non-response are corrected for using longitudinal row-and-column imputation procedures (see Little and Su 1989) and purely cross-sectional imputation techniques if longitudinal information is lacking. Thus, at least potential biases arising from the aforementioned selectivity can be reduced (see Frick and Grabka 2005).<sup>iii</sup>

## **2.2 Incidence and selectivity of PUNR**

### **2.2.1 Incidence of PUNR at person vs. household level**

There are at least two ways to express the overall incidence of PUNR in a household panel survey. For example, while “only” 5.4% of all adult household members did not fill in the requested questionnaire themselves in 2005, their non-participation behavior affected the measurement of relevant outcomes for all other members of their respective households as well. Thus 11.5% of all individuals (including children who not yet reached the respondent age of 17) lived in a household that was affected by PUNR.

Figure 1 presents time series information on the incidence of PUNR in the SOEP data over the period 1984 through 2007. While there was almost no PUNR in the starting wave 1984 (the contract with the fieldwork agency stated explicitly that in the first wave, households had to be “completely” interviewed with only a few strict exceptions allowed), there has been a clear tendency towards a growing percentage of non-participating respondents since then. This process is even more striking given the secular trend towards smaller households: the population at risk of PUNR is actually shrinking due to the increasing share of singles and lone parents in the population (with minor children up to 17 years of age, the respondent age in SOEP), i.e., household types for which non-response of the only respondent living in that household by definition yields a complete dropout (= UNR, unit non-response). Of course, the increase in the incidence of PUNR is also driven by the accumulation of “old” PUNRs over



time, i.e., persons who basically never give personal interviews while other household members continue to do so.

### **2.2.2 Incidence of PUNR by household size**

Following from this, a straightforward way to identify potential selectivity in PUNR comes with the number of adults (= target respondents) living in a given household. Obviously, those households with only one respondent do not bear the risk of PUNR. The risk of the household unit being affected by PUNR, however, clearly increases with the number of respondents; thus a household of six is more likely to be affected by PUNR than a household of only two adults. While children below respondent age do not bear the risk of PUNR themselves, they may be affected by PUNR of adult household members: any misrepresentation of the adults' resources in the household aggregate will affect measures of child poverty.

Figure 2 clearly illustrates this effect by indicating consistently increasing shares of individuals affected by PUNR, either directly (due to their own non-participation) or indirectly (due to non-responding co-residents): in 2005, for example, the share of individuals affected by PUNR within the household context is around 10% in households with two persons of respondent age, about 17% in households with three respondents, and above 20% for the rather few observations in large households of four and more adults.

### **2.2.3 Incidence of PUNR by panel experience**

For a long-running household panel such as SOEP, any non-response behavior is crucial for maintaining the quality and representativeness of the *longitudinal* data. The research on the scope and selectivity of UNR in household panel studies provides rather robust evidence that the probability of dropping out decreases with panel experience, thus, any additional interviews reduce the probability of UNR (see, e.g., Watson & Wooden 2009 on the Australian HILDA Survey). Although there is clear empirical support for this hypothesis on UNR in the SOEP data (on the weighting scheme in SOEP, see Kroh & Spieß 2007), the probability of PUNR does not necessarily monotonically decrease for long-term respondents (see Figure 3). We find a clear reduction in this probability only over the first few years of panel experience (including the years of childhood that a person spent “growing up” in the survey without being a respondent him/herself) and again after approximately 20 years.

It may be that a respondent in an otherwise cooperative household is simply more likely to drop out temporarily in a period of twenty years than in just five years. As such, a

temporary drop-out of a *long-term* respondent may be less problematic than that of a short-term respondent, i.e., it may reduce the number of observations in a balanced panel sample, thus reducing efficiency. However, an alternative hypothesis that is not tested in this paper could be that the established relationship between interviewer and respondent for long-term panel members makes these individuals more likely not to participate in times when “unusual” events occur, that is, occurrences that they are embarrassed to tell the interviewer about (e.g., a successful manager’s job loss). This second hypothesis is more in line with findings by Kapteyn et al. (2006), who argue that attritors in the Health and Retirement Survey (HRS) who are recruited back into the survey are very different from permanent attritors.

#### **2.2.4 Selectivity of PUNR**

To control for potential selectivity of PUNR, we make use of multivariate analyses. Table 1 shows results from a pooled probit regression model based on more than 45,000 individuals observed in the SOEP during the period 1985 to 2007. This totals more than 325,000 person-year observations; thus, we use robust standard errors obtained from clustering at the level of individuals.

Women tend to have a lower probability of PUNR, which is also true for the middle-aged (25-40) and elderly (66 years and over). Similarly, home ownership, an increasing number of dependent children, and increasing levels of education are negatively related to PUNR. Compared to the head of household, we find spouses or partners, children, and any other household members to be more likely to show PUNR. The probability of PUNR is particularly high at the time of a given household’s first interview and increases when changes in household composition occur. As expected, the risk of individual non-participation increases with the number of adults to be interviewed. We also find significant evidence that item non-response (INR) on the question dealing with current monthly household income is related to PUNR of at least one household member. This is also in line with findings based on data from various panel surveys showing a positive impact of INR on income questions in wave  $t$  on the probability of UNR (i.e., attrition) in wave  $t+1$  (see e.g., Loosveldt et al. 2002, Frick & Grabka 2010). Summing up, these findings may be taken as indications that economically active household members are more common among PUNR and thus probably major contributions to overall household resources are understated.

### 3 Dealing with PUNR in income-based analyses of economic well-being

Keeping the aforementioned selection issues in mind, the following section starts by briefly introducing the various approaches to dealing with PUNR before demonstrating prototypical empirical applications from a welfare economics perspective using an aggregated measure of equivalent household income.

#### 3.1 Alternative approaches

There exist a variety of ways to deal with PUNR in empirical analyses, four of which will be applied in this paper: *Ignorance*, i.e., assuming the missing individual's income to be zero. (b) *Adjustment* of the equivalence scale to account for differences in household size and composition. (c) *Elimination* of all households observed with PUNR using subsequent re-weighting procedures. (d) *Longitudinal Imputation* of the missing income components.

- Ignorance: Assume the individual affected by PUNR has no income of their own to add to the household's overall resources, but she does have needs that ought to be considered when constructing the household's equivalence scale. This means effectively ignoring PUNR in the measure of household *income*, while continuing to consider her *needs*, i.e.,  $Y(PUNR)=0$  &  $Needs(PUNR)>0$ .
- Adjustment: Assume the individual has no income of her own as well as no needs, thus completely ignore the existence of the individual with PUNR. This effectively means deleting non-responding individuals from PUNR households by adjusting the respective equivalent scale downward, i.e.,  $Y(PUNR)=0$  &  $Needs(PUNR)=0$ , which implies that income and needs of the missing individual are identical to those of the observed household members.
- Elimination: Delete *all* individuals living in households affected by PUNR (i.e., also the successfully interviewed persons) and rescale the population weights for those households that bear a risk of PUNR but did fully complete the survey. This assumes that the income and needs of households with PUNR are mirrored by successfully completed households with two and more respondents.<sup>iv</sup>
- Imputation: Impute any income measure missing due to PUNR, thus considering all households with completed information on income as well as needs by assigning incomes

to PUNRs on the basis of comprehensive (cross-sectional and longitudinal) imputation procedures (details are given in the next section below).

- Finally, another approach that will not be considered in the remainder of this paper: the “flat correction factor,” which has been applied in the European Community Household Panel (ECHP) (see Eurostat 2000).<sup>v</sup> In this case, each household is assigned a specific “within-household non-response inflation factor” (see Appendix A for details). The basic assumption underlying this approach is that all income components in a given household are affected by PUNR in the very same way—thus, even if this were considered a pseudo-correction of the misreported income *level*, the income *portfolio* of the household would most likely remain subject to bias.

## 3.2 The imputation strategy

The imputation procedure correcting for income missing due to PUNR is based on the following principles: impute the most detailed income information possible (i.e., use different components of income), make use of household context data and, most important from the panel perspective, employ longitudinal information if available.

### 3.2.1 Imputation of single income categories

As is standard in the welfare economics literature, our target income measure is annual post-government income of the previous year, which is given by the sum of all market incomes (from labor, capital, and private transfers) *plus* pensions and public and private transfers received *minus* taxes and social security contributions, aggregated across all household members.<sup>vi</sup> However, instead of imputing just a “lump sum” of missing income, we aim at imputing six individual gross income components that are directly compatible with the more detailed information collected in the standard SOEP questionnaire every year. This allows aggregation at the household as well as the tax unit level<sup>vii</sup> to match the very same income aggregates for all individual observations, whether PUNR or successfully interviewed. This aids in the final simulation of direct taxes and social security contributions, which explicitly requires considering the interdependence of income and tax calculations for joint tax filers. In so doing, we can derive a consistent measure of “Household Post-Government Income” as the major source for inequality analyses. Finally, this procedure is likely to exert less bias in portfolio analyses than a “lump sum” or “flat factor” approach would.

For each PUNR, we impute the following six income components, which are collected at the individual level in the SOEP (all other income components such as means-tested public transfers or capital income are surveyed at the household level, and thus, by definition, are already included in the income measures derived from the successfully interviewed household members):

1. labor income (the sum of all incomes from dependent employment, self-employment, secondary jobs, including extra pay such as Christmas or discretionary bonuses, etc.),
2. pensions (the sum of all pensions received from the statutory social security pension system (GRV) as well as from the tax-financed pension system for civil servants, including any survivor benefits in both systems),
3. unemployment compensation (the sum of assistance received through the unemployment insurance scheme, unemployment assistance, and subsistence allowance from the labor office),
4. (public) student aid,
5. (public) maternity leave benefits, and
6. private transfers (including alimony).

### **3.2.2 Imputation of income received and amount received**

For each of the components mentioned above, we employ a two-step imputation procedure. First, we need to define a “filter,” indicating whether a given person received the respective component, and conditional on predicted receipt, we need to impute a positive value for that income.<sup>viii</sup> In both cases, we make use of longitudinal information, if at all available, which has been shown to clearly improve the quality of imputation (see, e.g., Spiess & Goebel 2003 using ECHP, Starick & Watson 2007 using HILDA, Frick & Grabka 2010 using SOEP, HILDA and BHPS data). We separate all observations with PUNR into four groups, depending on the availability of information from the previous or following year, from both years, or from neither. In fact, for any PUNR with missing information in year  $t$  and a successful interview in  $t-1$ , we can derive valuable and highly predictive information for the target information in  $t$  from previous year’s income receipt at the time of the interview when she was asked about her current income and employment status.

As such, receipt of a given income  $Y_k > 0$  ( $k$  = six income components) is predicted on the basis of a multivariate probit model estimating the probability of income receipt in the observed population. For observations without any longitudinal information, we employ only contemporary control variables including individual information on sex, age, relationship to

the household head<sup>ix</sup>, and a range of household context information<sup>x</sup>. For observations with such longitudinal data, we also include income and employment status from the adjacent waves. In any case, if the predicted probability for receipt of a given income exceeds a randomly chosen threshold (drawn from a normal distribution with mean 0.5 and standard deviation 0.2 to consider uncertainty of the underlying imputation process), we assign a value of one to the respective filter indicator. In order to adequately control for the interdependency of income receipt, we include predicted filter dummies in subsequent regressions in the following order. We first estimate the probability for receipt of pensions, then include the predicted pension dummy in the estimation for the receipt of unemployment assistance. Both of these are then considered when running the probit regressions for maternity leave benefits and student grants. All four filter dummies are then included in the regression of received private transfers, and finally, we use all five filter variables to predict receipt of labor income. It should be noted that the predictive power of these models is very high, especially when introducing longitudinal data: e.g., the pseudo- $R^2$  for estimating the probability of receiving labor income ranges up to 0.7.

In a second step, we predict the amount of income conditional on predicted receipt as indicated in the filter variables. Here, we again make use of longitudinal data in the imputation process by applying the row-and-column imputation as described in Little & Su (1989).<sup>xi</sup> Longitudinal information is used over a seven-year (or wave) shifting window around the point in time with missing data—up to three years before and after the occurrence of PUNR. Assuming that information obtained from observations more distant from the missing data point is less strongly correlated to the missing information, we assign decreasing weights to more widely separated information.<sup>xii</sup>

The row-and-column procedure proposed by Little & Su<sup>xiii</sup> is carried out as follows: The column effect  $c_t$  is defined by the relative cross-sectional annual income for each of the seven ( $k$ ) observations, thus capturing simple period effects:  $c_t = Y_t / \frac{1}{k} \sum_{i=1}^k Y_i$  (with  $Y_t$  denoting total incomes at time  $t$  and  $k$  being seven years). Second, the row effect  $r_i$  is based on the longitudinal income data collected from a given individual  $i$ ; thus, it gives the within-person average income position  $r_i = \frac{1}{k} \sum_{j=1}^k \frac{Y_{ij}}{Y_j}$  (with  $k$  denoting the number of valid income values for individual  $i$  and  $Y_j$  denoting the average income at time  $j$ ).<sup>xiv</sup> Multiplying the column effect  $c_t$  with the row effect  $r_i$  yields the expected income position of individual  $i$  at time  $t$ :  $E_{it} = r_i * c_t$ .

Finally, a stochastic component is introduced by considering the deviation of the expected and the observed value from the nearest neighbor. Combining those three effects yields an estimate for the missing income  $Z_{it} = E_{it} * [Y_{nt} / E_{nt}]$  (with  $Y_{nt}$  and  $E_{nt}$  denoting the observed and the expected income of the individual  $n$ , defined as the nearest neighbor of individual  $i$ ).

This procedure provides imputed values for all PUNRs with at least one valid interview within the seven-year window under investigation; however, it fails to do so if no longitudinal data is available. This applies to less than one-third of all PUNRs and thus requires the application of a purely cross-sectional imputation strategy for the various income components, some of which are closely linked to the life course, such as student aid, maternity benefits, and pensions. Conditional on the predicted receipt we impute the money metric value for the six income components, partly also separately by gender. Again, in order to control for eventual interdependence in receipt of the various income sources, the list of regressors includes the full set of dummies for receipt of the other (five) components, as well as the same set of variables used to predict the filter information. In order to preserve variation when predicting the respective income value for PUNRs, we again introduce a stochastic component drawn randomly from the residuals of the regression sample.

### 3.2.3 Results of the imputation process

A straightforward assessment of the overall impact of these imputation procedures is given in Figures 4a-c presenting time series information on a comparison of observed and fully imputed values for the various income components. We show (a) the population share holding a given component, (b) mean values for each component (in nominal euros) conditional on receipt, and (c) the resulting mean values for the entire population.

According to Figure 4a and in line with the regression results on the selectivity of PUNR presented above (Table 1), non-participating individuals are being assigned labor incomes clearly more often than is true among the observed population. In fact, our imputation procedures impute labor incomes for roughly 75% of all individuals with PUNR while only 60% to 70% of the observed individuals report receipt of labor income throughout the previous year. Accordingly, the share of individuals for whom we imputed receipt of any other income components (pensions, unemployment benefits, maternity benefits, student grants, private transfers) is clearly below the level among the successfully interviewed population.

Comparing the levels of income for those who have been either observed or imputed as recipients of a given income component (see Figure 4b), we differentiate between indi-

viduals where we could apply the longitudinal imputation according to Little & Su (1989) and those where a purely cross-sectional approach was necessary due to lacking longitudinal information. By and large, for all income components, both types of imputation yield similar average values—except for labor income, where we can identify a consistently higher average income among the longitudinally imputed individuals. This may partly reflect a statistical artifact since for new entrants into the labor market, who typically have rather low earnings, we do not observe any previous income receipt of a type that requires applying the L&S imputation procedure.

Finally, Figure 4c gives the average values for the various components across the entire population before and after full imputation. While the general result appears to be that our imputation does not alter the values very much, one should keep in mind that these figures also pertain to all fellow household members affected by PUNR due to the pooling and sharing assumption underlying the calculation of an equivalent household income.

Thus, before carrying out the welfare analyses, we need to incorporate the relevant imputed gross income components into the simulation of direct taxes and social security contributions, then add any public transfers, and finally recalculate a PUNR-adjusted measure of post-government income.

### 3.3 Intermediate conclusions and hypotheses

What impact might the choice of one of the alternative treatments (*ignorance, adjustment, elimination, imputation*) exert on results of inequality and poverty analyses based on microdata affected by PUNR? First of all, given the secular trend toward an increasing incidence of PUNR over time (see Figure 1 above), one should expect any bias (in measures of inequality, poverty, intra-household distribution, and aggregates) arising from PUNR to be increasing over time, i.e., with the duration of the panel. Secondly, the selectivity of PUNR as shown above, clearly challenges the basic assumption underlying the various approaches: Version 1 “ignoring PUNR” [thus assuming  $Y(PUNR)=0$ ] appears the least plausible given that PUNR is more prominent among economically active household members. In principle, this critique also applies to the basic assumption underlying version 2, “adjusting equivalence scale,” where we assume the incomes and needs of the non-participating members to mirror those of their fellow household members who were interviewed [ $Y(PUNRs) \sim Y(noPUNRs)$  within PUNR-HH]. Finally, there may be no clear-cut answer about distributional effects arising from choosing version 3 “elimination and re-weighting” [i.e., incomes of PUNR



households are equal to those of non-PUNR households,  $Y(PUNR-HH) \sim Y(noPUNR-HH)$ ] or version 4 “imputation of PUNR” [ $Y(PUNR)=f(X)+e$ ]. It can be assumed, however, that the more similar the variables used in the re-weighting scheme and in the imputation procedure are, the more similar the treatment effects on inequality and poverty will be. In any case, while both approaches appear clearly less normative than approaches 1 and 2 due to (adequately) controlling for selectivity, the imputation approach should be preferred because it retains the complete panel population, which is especially important for mobility research.

## 4 Empirical analyses: inequality and mobility effects

The following empirical analyses are based on annual post-government income received in the calendar year preceding the survey year (including a measure of net imputed rent, see Frick & Grabka 2003). For cross-temporal comparability, we express all incomes in 2000 prices, also correcting for regional purchasing power differences between East and West Germany up to the mid-1990s. To adjust for different income needs across households due to differences in size and age composition, we apply the modified OECD equivalence scale, assuming adult household members (above the age of 14) to have 50% of an adult’s income needs and children up to age 14 to have 30% of those needs. In the following, we will compare results obtained from the four approaches to deal with PUNR on measures of income inequality, poverty and mobility. With respect to poverty measures, we will also try to identify the degree to which results do not only coincide for the entire population, but we will also look at the consistency of those alternatively derived measures for each individual. Here, it will not only be important to find out whether the two approaches yield, e.g., a similar share of individuals at risk of relative income poverty, but also whether these results are identical for the very same persons.

In order to provide a more robust picture, we apply a range of established indicators used in the literature. We measure income inequality by means of the Gini coefficient, the mean log deviation (MLD), which is more sensitive to changes at the lower end of the income distribution, and the top-sensitive half-squared coefficient of variation (HSCV). Relative income poverty will be measured based on a poverty threshold of 60% of the median. In order to also identify eventual effects *within* the population defined as poor, we make use of the family of poverty measures developed by Foster, Greer, & Thorbecke (1984), allowing the poverty aversion parameter alpha to take on the values of 0 (poverty risk rate), 1 (normalized poverty gap), and 2 (giving higher weight to those further below the poverty threshold). Fi-

nally and particularly important for addressing the relevance of these alternative treatments for panel research, income mobility is measured on the basis of the measures introduced by Fields & Ok (1999) and by Shorrocks (1978).<sup>xv</sup>

## 4.1 Inequality and Poverty

### 4.1.1 Hypothetical effects

What are the hypothetical effects of treating PUNR, in whatever way, on income levels, relative poverty, and inequality? First of all, any explicit accounting for PUNR should yield *higher* incomes among those affected by PUNR; thus, average incomes (mean and median) of the entire population should also be subject to increase. Following from that, we should expect an increase in the relative poverty threshold and consequentially, the relative poverty risk among households *not* affected by PUNR (and thus without a change in their PUNR-adjusted income measure) should be higher than without PUNR-correction, other things being equal. Version 2 (*adjustment*) and Version 4 (*imputation*) explicitly yield higher equivalent income among households affected by PUNR. Thus, for this group, one should expect, *ceteris paribus*, decreasing poverty risk rates as long as their increase in household income exceeds the increase in the national poverty threshold.

In light of these two contradictory effects, the overall (net) effect of PUNR treatment on relative poverty risks at a *given point in time* remains unclear. However, due to the secular increase in the incidence of PUNR over time, *poverty trends* might be affected as well. Most likely, however, there will be effects on the *socio-demographic structure* of poor households. Almost by definition, the increase in the poverty threshold will cause an increase in poverty among those households not at risk of PUNR (i.e., single adults and lone parent families with only one household member of respondent age). Similar effects may be expected for households that bear the risk of PUNR but that were nevertheless completely interviewed.

### 4.1.2 How does PUNR affect income levels, inequality, and poverty?

Putting numbers to those considerations, we now turn to a comparison of results from inequality analyses based on the four approaches (labeled *Version 1*: “Ignoring”, *Version 2*: “Adjusting needs”, *Version 3*: “Deleting & re-scaling” and *Version 4*: “Imputation”). Figure 5 gives time series information on various percentiles (P10, P25, P50=Median, P75, and P90) of the respective annual equivalent post-government income. Apparently, in the early waves of

the panel, when PUNR was a rather rare event, there was almost no variation across our four measures. However, in line with the increasing incidence of PUNR over time (see Figure 1), we observe a clear and consistent differentiation: across the entire distribution, the results for Version 1 (ignoring) are, as speculated above, lower than in the other three treatments. These differences clearly pick up over the course of time. Adjusting the equivalence scale (Version 2), which implicitly means that incomes and needs within PUNR-households are correctly specified by the observed individuals, also controls only insufficiently for the underlying selectivity, whereas Version 3 (deleting and re-scaling weights) and Version 4 (imputation) yield the highest income levels—while being very similar in scope. As can be expected from the increasing deviation between top and bottom income levels in Figure 5, there is a secular trend toward increasing income inequality in Germany, which has accelerated substantially since the turn of the millennium (see Grabka & Frick 2008).

Figure 6 confirms this finding using various inequality indicators.<sup>xvi</sup> More importantly for our argument, however, we observe again an increasing gap among the four treatments, with Version 1 showing the highest degree of inequality (no matter which indicator is chosen), Version 2 yielding a somewhat lower degree of inequality, and finally Versions 3 and 4 producing—once again in similar fashion—the lowest level of dispersion.

Finally, we present results on relative income poverty using the parametric family of FGT measures. Confirming the hypotheses laid out in Section 4.1.1, the results for the poverty head count ratio (FGT0) in Figure 7 are highest for the Version 1 (“ignoring”), while Version 2 takes on a middle position, and “reweighting” as well as “imputation” yield the lowest level of poverty risk. Again the deviation is growing with duration of the panel, thus also reflecting the increase in PUNR incidence. The difference in poverty risk rates in the most recent years is up to two percentage points!

All these results are very stable using higher poverty aversion parameters in FGT1 (“normalized income gap ratio”) and FGT2 (which gives more weight to the poorest poor).

#### **4.1.3 Consistency of poverty “assignment” when using different approaches to tackle PUNR**

Obviously, there is clear variation in the overall level of relative poverty at the aggregate or national level across the various techniques. However, even if the results were more similar, it would not necessarily require that the various approaches identified the very same individuals as being poor. The results presented in Figure 8a and Figure 8b challenge the

consistency of the poverty measurement on the basis of the four approaches, thus answering the question of “who is poor according to approach x, but not poor according to approach y” and vice versa.

In other words, above and beyond the sheer interest in the overall share of people living below the poverty line, it is of utmost importance for the design, application, and evaluation of social policy programs to know more about the socio-economic structure of the population affected by relative income poverty. The effects of any such reform should certainly not just mirror the assumptions underlying the approach to deal with PUNR in the microdata used. If that were the case, for example, we would expect child poverty to look different if PUNR were mostly a problem among households with dependent children, thus misreporting their income and making them appear poorer than they actually are.

In Figure 8a, we restrict our sample to those who live below the poverty line according to Version 4 (“imputation”). The time series graphs show the share of those identified as *non-poor* according to the three other approaches. In line with the analysis results so far, there is a high degree of concordance with the identification of poverty in Version 3 (“elimination and re-weighting” as given by the yellow graph), although the population used in Version 3 is more selective, as illustrated by the share of those missing from the analysis (light blue graph) which accrues up to about 10% of the baseline population. Clearly less comparable are the results based on Version 1 (dark blue graph) and Version 2 (pink graph), which both show an increasing deviation from the results obtained from the imputed data.

In contrast, Figure 8b is based on the population of *non-poor* in Version 4 (“Imputation”); and the various graphs show time series of the share of those persons identified as poor in Versions 1 to 3. Again, we observe a high degree of similarity with the results of Version 3, while the share of those eliminated in Version 3 is as high as 12% in the most recent waves. The results obtained from Versions 1 and 2 are significantly less comparable, and differences again are growing over time.

## 4.2 Poverty and Income mobility

The analyses so far indicated a significant impact of the methodological decision on how to cope with PUNR on cross-sectional results (inequality and poverty). In the following we address the question of the degree to which this is true for longitudinal analyses as well by comparing results derived from the four methods with respect to poverty and income mobility.

Figure 9 presents results from simple wave-to-wave poverty mobility analyses, for robustness purposes averaging all pooled two-year balanced panels over the period 1985-2007. For each of the four approaches, we show the share of individuals moving into or out of poverty within a given two-year interval. In order to better assess the impact of PUNR on poverty mobility, we separate the population into three groups: those who were continuously living in households not affected by PUNR in both waves, those who experienced PUNR in only one wave, and finally the group of people affected by PUNR in both years. Comparing the four techniques, the selectivity of the population in Version 3 (elimination and re-scaling) becomes apparent. Indeed, one may argue that maintaining the entire survey population, i.e., including households with PUNR, may be more important for mobility analyses than for purely cross-sectional analyses, a strong argument in favor of imputation (Version 4) over weighting (Version 3).

Although the degree of poverty mobility in the aggregate does not differ much across the four versions, PUNR households show much higher mobility rates, simply because PUNR in at least one wave increases the probability of being poor due to understated incomes as shown above in Section 4.1. Results based on imputed data (Version 4) still exhibit above-average mobility, in particular if PUNR was present only in one wave. On the one hand, this partly reflects the uncertainty of the underlying imputation procedure (i.e., we inject variation by adding residuals to the predicted incomes); on the other hand, we cannot rule out that the missingness simply reflects “true” mobility, if the PUNR, for example, had been caused by a change in labor market status of the respondent that interfered with survey participation. If the latter was the case, then indeed the mobility results in Version 3 (elimination and re-scaling) would be downward-biased.

Using the same data, Figure 10 gives very consistent results with respect to income mobility over two years when applying the Fields & Ok (1999) index. It should be noted that the results presented here are insensitive to the choice of the mobility measure; applying, for example, the Shorrocks (1978) index (using the Gini coefficient) yields more or less identical results (available from the authors upon request).<sup>xvii</sup>

## 5 Conclusions

Using 24 waves of panel data from the German SOEP, we find an increasing incidence of PUNR together with clear indications of the selectivity of PUNR. A major consequence of this phenomenon is a systematic downward bias in the level and development of income inequality and relative poverty, whereas income mobility will be overstated due to people moving into and out of PUNR. Our strategy of imputing single income components as well as including them in the estimation of taxes and social security contributions appears to be an appropriate means to cope with both of these problems, thus guaranteeing a less biased measure of post-government income as the empirical basis for analyses of income inequality and poverty.<sup>xviii</sup> The imputation of various components instead of just adjusting the “annual post-government income measure” (e.g., by means of a “flat correction factor”) may also be considered advantageous because it supports decomposition analyses (by income source) and portfolio analyses while maintaining the entire survey population (in contrast to alternative strategies that exclude those affected by PUNR and re-weight those at risk of PUNR).

Future research will have to address the following questions:

- Is there a correlation between PUNR und subsequent attrition (UNR) that would be relevant for mobility analyses?
- Does the choice of PUNR treatment affect comparability in cross-country comparisons (see Frick & Grabka 2010 for the need to harmonize the procedures used for the imputation of item-non-response)?
- While our analysis dealt with the missing contribution of individuals to their respective household’s resources, PUNR may also yield a similar bias in other research areas, such as labor economics, where the interaction among household members is of crucial importance—for example, when modeling labor supply decisions of couples. The missing mechanism for PUNR may not be random at all if the individual is unable to participate in the survey because, for example, she is earning a great deal of money drilling for oil on an offshore platform, or simply because she does not want to answer questions while unemployed or severely ill.
- Finally, what do our results imply for designing incentives targeted at increased participation in household panel surveys (see Laurie & Lynn 2009, DeLeeuw et al. 2003, Hill & Willis 2001)? Instead of correcting the microdata *after* data collection, one should instead try to prevent missing data from occurring. Thus, it might be preferential from a data col-

lection point of view to consider ideas for early interventions, such as (a) collecting proxy information on individuals with restricted interview capability, e.g., due to severe sickness, dementia, Alzheimer's, etc.; (b) increasing incentives to participate, e.g., through monetary incentives. While this additional participation may exert positive spillover effects on other household members, it may also yield some habituation effect from a panel perspective: interviewees who were paid extra money once may want to keep their "bonus" in future waves as well, making this approach a rather expensive one. One also can only consider providing an additional incentive at the household level if all adult respondents participate (this is done in the HILDA survey; see Watson & Wooden 2009). (c) An alternative might arise with a short drop-off questionnaire to be filled in by PUNR respondents, in order to improve the basis for the imputation or weighting procedure. However, as is the case for proxy interviews, such an approach might also be used by respondents to "sneak out" of the regular survey in order to reduce the response burden.

Above and beyond the arguments brought forward in this paper, there may also be other reasons why PUNR is increasing in prevalence over the last few years than simply measurement issues. One argument arises from the increasing number of individuals who have *multiple residences*, e.g., due to either long-distance commuting between "home" and "work" and to choosing "modern" lifestyles such *living apart together* (Asendorpf 2009); that is, couples who do not share a common address but rather live in two separate places. By making it difficult to determine which household a given person belongs to and whether her resources and needs should be assigned to just one or (partially) to several households, these recent social developments make it ever more difficult to precisely define the concept of "private households."

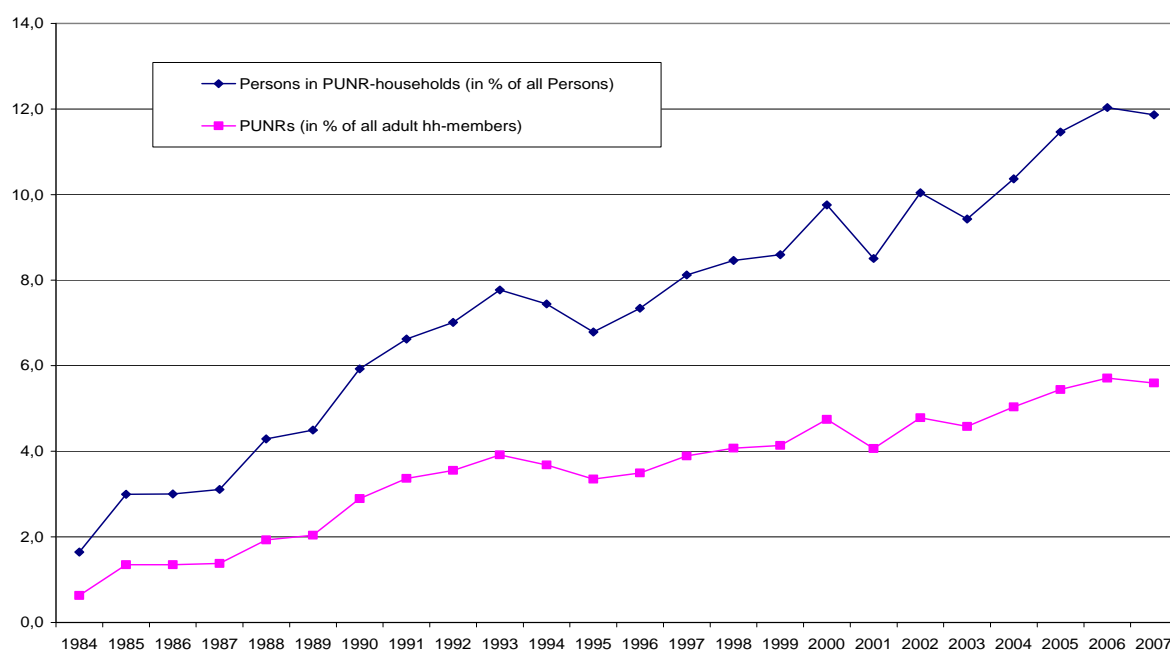
## 6 References

- Asendorpf, J. B. (2009): Living Apart Together: Alters- und Kohortenabhängigkeit einer heterogenen Lebensform. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 60(4): 749-764.
- Atkinson, A. B. and F. Bourguignon (eds) (2000): *Handbook of Income Distribution*, North-Holland.
- Barceló, C. (2008): The impact of alternative imputation methods on the measurement of income and wealth: evidence from the Spanish survey of household finances, Documentos de Trabajo N.º 0829, Banco De Espana, Madrid.
- Becker, G. S. (1991): *A Treatise on the Family*, Enl. Edition, Cambridge University Press.
- Bonke, J. and Uldall-Poulsen, H. (2007): Why Do Families Actually Pool Their Income? Evidence from Denmark, *Review of Economics of the Household*, 5(2): 113-128
- Buhmann B, Rainwater L, Schmaus G, and Smeeding T. M. (1988): Equivalence scales, well-being, inequality, and poverty: Sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) database. *The Review of Income and Wealth*. 34(2):115–142.
- Canberra Group (Expert Group on Household Income Statistics) (2001): Final Report and Recommendations. Ottawa, <http://www.lisproject.org/links/canberra/finalreport.pdf>.
- De Leeuw E. D., Hox J. J., Huisman M. (2003): Prevention and treatment of item nonresponse. *Journal of Official Statistics*, 19, 153-176.
- De Leeuw, E. D., Hox, J. J., Dillmann, D. A. (eds)(2008): *International Handbook of Survey Methodology*. New York, London: Erlbaum, Taylor & Francis.
- Eurostat (2000): Imputation of Income in the ECHP. DocPan 164/00. Luxembourg.
- Fields, G. S. and Ok, E. (1999): Measuring Movement of Incomes, *Economica*, 66: 455-471.
- Foster, J., Greer, J. and Thorbecke, E. (1984): A Class of Decomposable Poverty Measures. *Econometrica*, 52 (3), 761-766.
- Frick, J. R. and Grabka, M. M. (2003): Imputed Rent and Income Inequality: A Decomposition Analysis for the U.K., West Germany, and the USA. *The Review of Income and Wealth*, 49(4): 513-537.
- Frick J. R. and Grabka M. M. (2005): Item non-response on income questions in panel surveys: incidence, imputation and the impact on inequality and mobility. *Allgemeines Statistisches Archiv*, 89, 49-61.
- Frick, J. R. and Grabka, M. M. (2010): Item Non-Response and Imputation of Annual Labor Income in Panel Surveys from a Cross-National Perspective, In: J. A. Harkness, B. Edwards, M. Braun, T. P. Johnson, L. E. Lyberg, P. Ph. Mohler, B.-E. Pennell, T. Smith (eds): *Survey Methods in Multicultural, Multinational, and Multiregional Contexts*, Wiley & Sons. (in press).
- Grabka, M. M. and J. R. Frick (2008): The Shrinking German Middle Class—Signs of Long-Term Polarization in Disposable Income? DIW Berlin Weekly Report No. 4/2008, 21-27.
- Groves, R. M. and Couper, M. P. (1998): *Non-response in Household Interview Surveys*. New York: Wiley.
- Groves, R. M., Dillman, D. A., Eltinge, J. L. and Little R. J. A. (eds.) (2002): *Survey Nonresponse*. New York: Wiley.
- Hill, D. and Willis, R. J. (2001): Reducing Panel Attrition: A Search for Effective Policy Instruments. *Journal of Human Resources*, 36, 416-438.
- Kapteyn, A., P.-C. Michaud, J. Smith and A. van Soest (2006): Effects of Attrition and Non-Response in the Health and Retirement Study, IZA Discussion Paper No. 2246, August 2006, IZA Bonn.
- Kroh, M. and M. Spieß (2007): Documentation of Sample Sizes and Panel Attrition in the German Socio-Economic Panel (SOEP) (1984-2007). Data Documentation 39, Berlin: DIW Berlin.
- Laurie, H. and Lynn, P. (2009): The use of respondent incentives on longitudinal surveys. In P Lynn (ed.), *Methodology of Longitudinal Surveys*. London: John Wiley and Sons, 205-234.



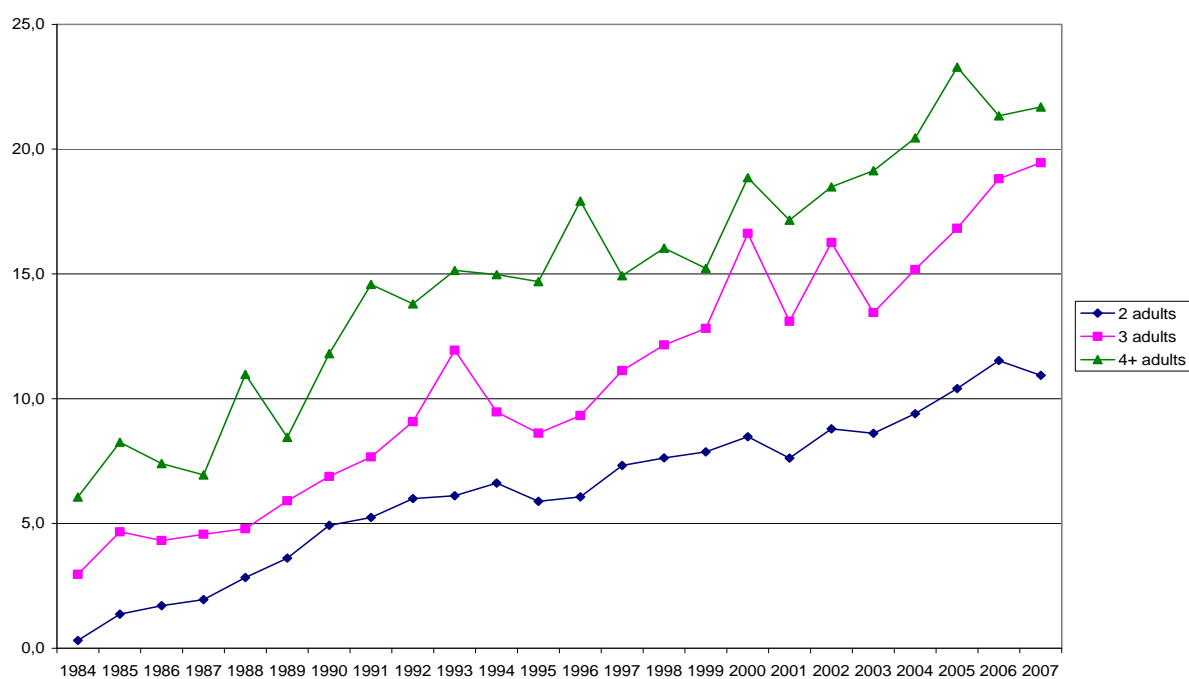
- Little, R. J. A. and Su, H. L. (1989): Item Non-Response in Panel Surveys. In D Kasprzyk, G Duncan, MP Singh (eds.), *Panel Surveys* (pp. 400-425). New York: John Wiley.
- Loosveldt, G., Pickery, J. and Billiet, J. (2002): Item non-response as a predictor of unit non-response in a panel survey. *Journal of Official Statistics*, 18, 545-557.
- Lynn, P. (ed.) (2009): *Methodology of Longitudinal Surveys*. London: John Wiley and Sons.
- Nicoletti, C. and Peracchi, F. (2006): The effects of income imputation on microanalyses: evidence for the European Community Household Panel. *Journal of the Royal Statistical Society Series A*, 169 (3), 625-646.
- Rubin, D. B. (1987): *Multiple imputation for nonresponse in surveys*. New York: Wiley.
- Shorrocks, A. F. (1978): Income Inequality and Income Mobility, *Journal of Economic Theory*, 19: 376-393.
- Smeeding, T. M. and D. H. Weinberg (2001): "Toward a Uniform Definition of Household Income," *The Review of Income and Wealth*, 47(1), 1-24.
- Spiess, M and Goebel, J. (2003): Evaluation of the ECHP imputation rules. CHINTEX Working Paper No. 17.
- Starick, R. and Watson, N. (2007): Evaluation of Alternative Income Imputation Methods for the HILDA Survey. HILDA Project Technical Paper Series No. 1/07, Melbourne Institute of Applied Economic and Social Research, University of Melbourne.
- Vogler, C. and J. Pahl (1994): Money, Marriage and Inequality within Marriage. *Sociological Review*, 42(2), 236-288
- Wagner, G. G., Frick, J. R., Schupp, J. (2007): The German Socio-Economic Panel Study (SOEP) – Evolution, Scope and Enhancements. *Schmoller's Jahrbuch – Journal of Applied Social Science Studies* 127(1): 139ff.
- Watson, N. and Wooden, M. (2009): Identifying factors affecting longitudinal survey response. In P Lynn (ed.), *Methodology of Longitudinal Surveys*. London: John Wiley and Sons, 157-182.

**Figure 1:** Incidence of PUNR in the German SOEP, 1984-2007



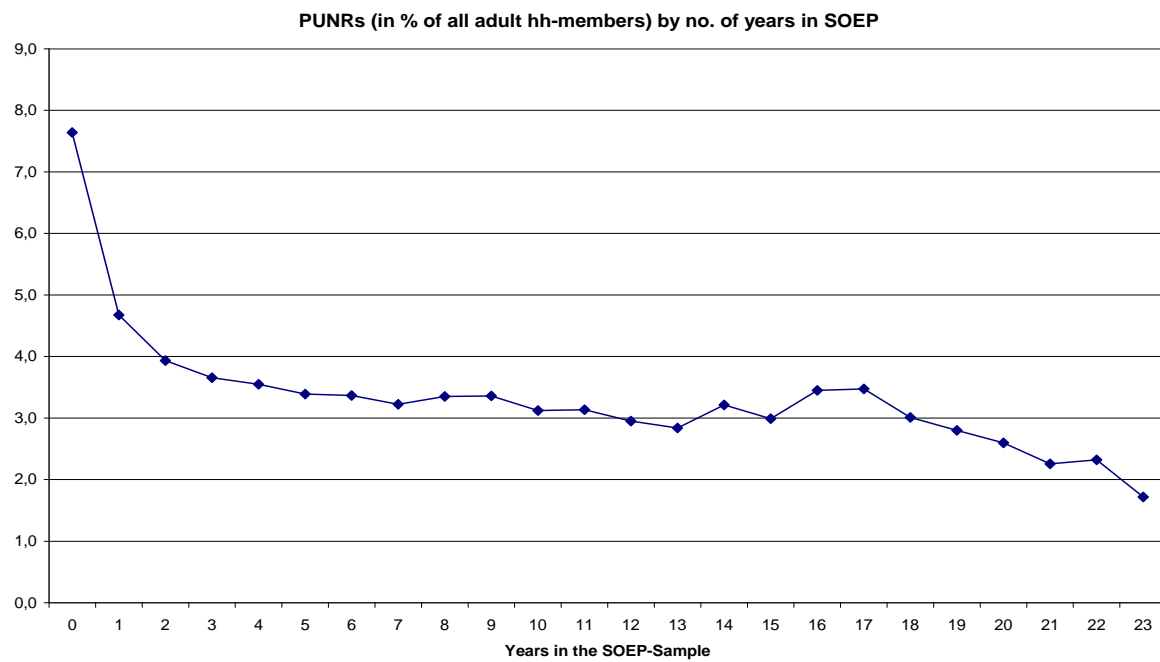
Source: SOEP. Authors' calculations.

**Figure 2:** Incidence of PUNR by number of adult household members



Source: SOEP. Authors' calculations.

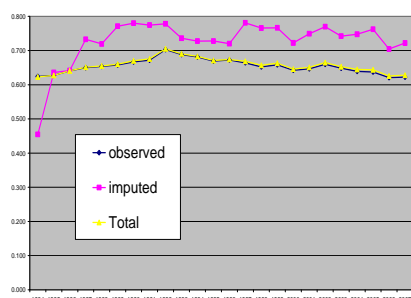
**Figure 3:** Incidence of PUNR by panel duration



Source: SOEP. Authors' calculations.

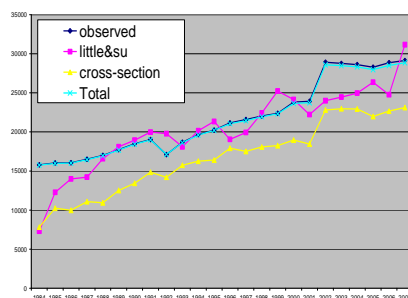
**Figures 4.a-c:** The impact of imputation: population share holding income components and average income

A. Share of persons with  $Y > 0$

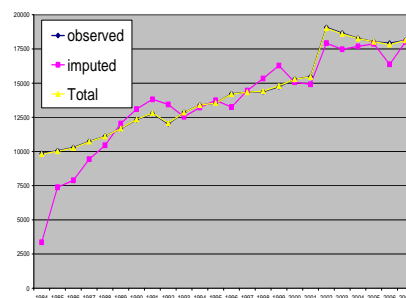


B. Mean values  $> 0$

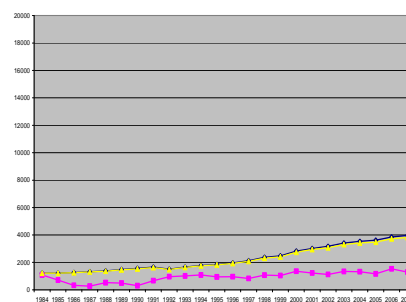
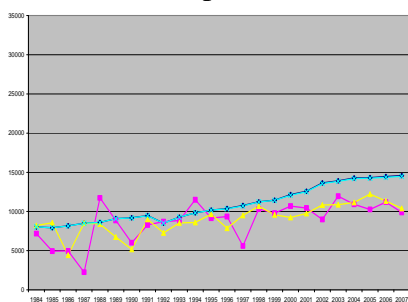
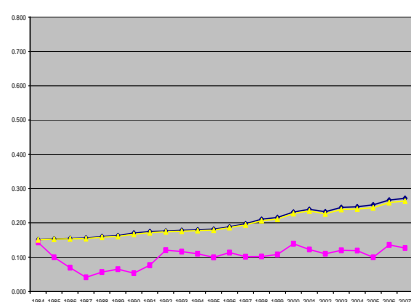
(1) labour incomes



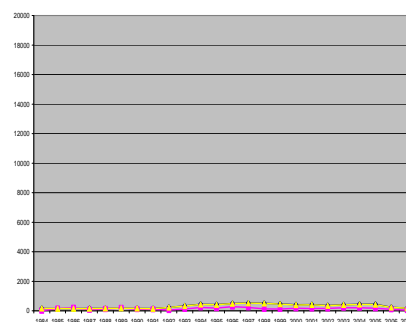
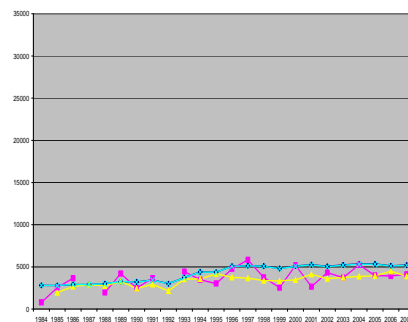
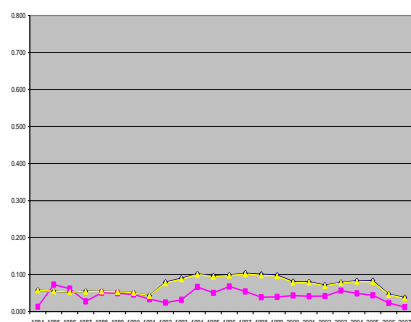
C. Mean values (incl. 0)



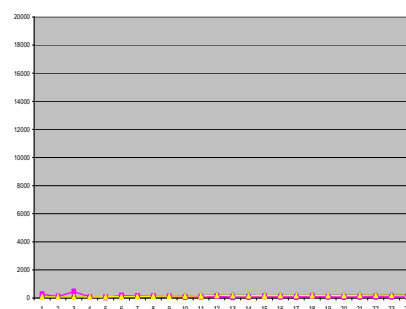
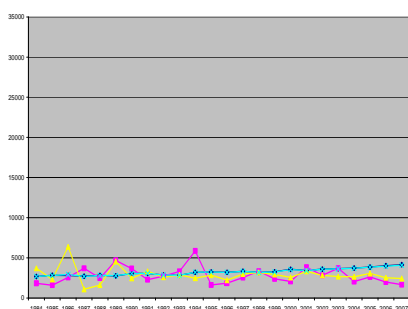
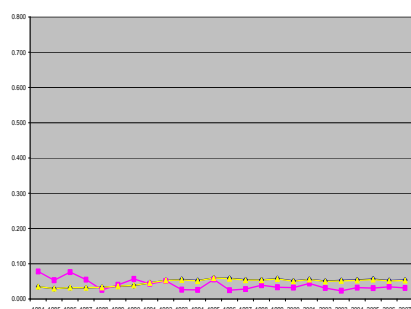
(2) pensions



(3) unemployment benefits

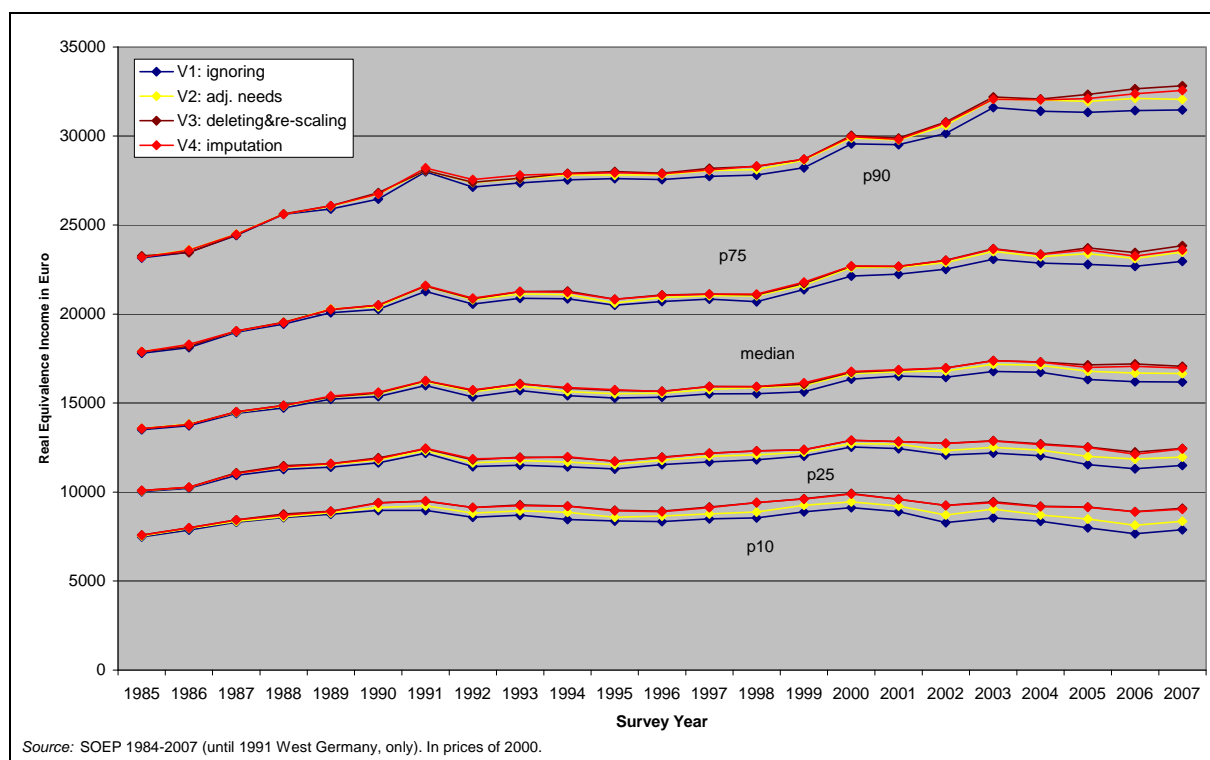


(4) other transfers (maternity benefits, student grants, private transfers)

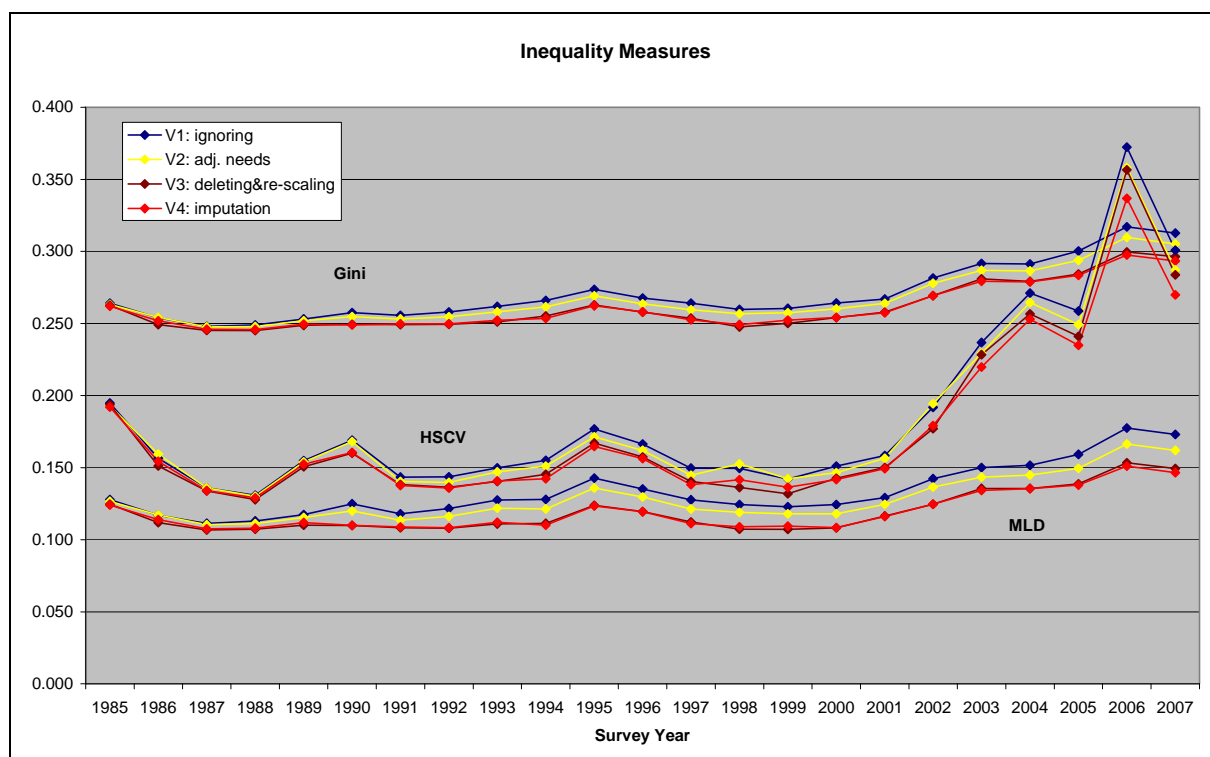


Source: SOEP. Authors' calculations.

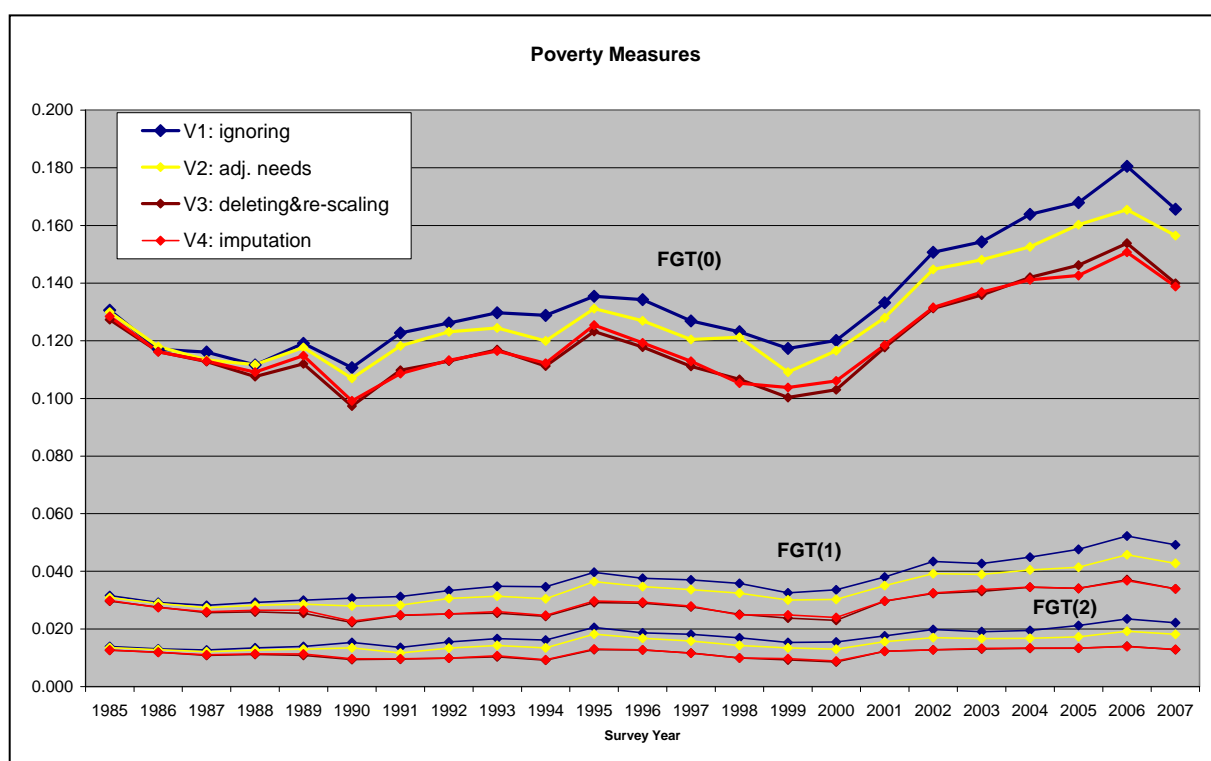
**Figure 5:** The impact of PUNR treatment on the distribution of equivalent income



**Figure 6:** The impact of PUNR treatment on income inequality

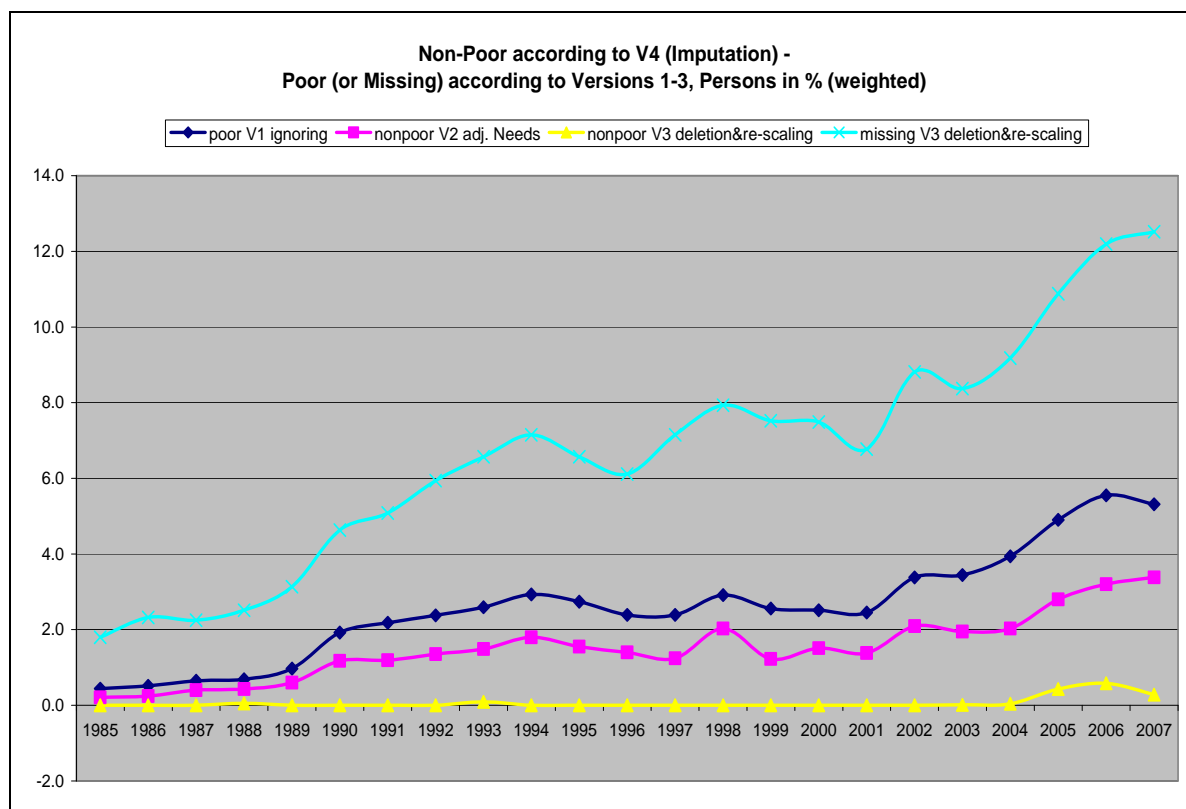
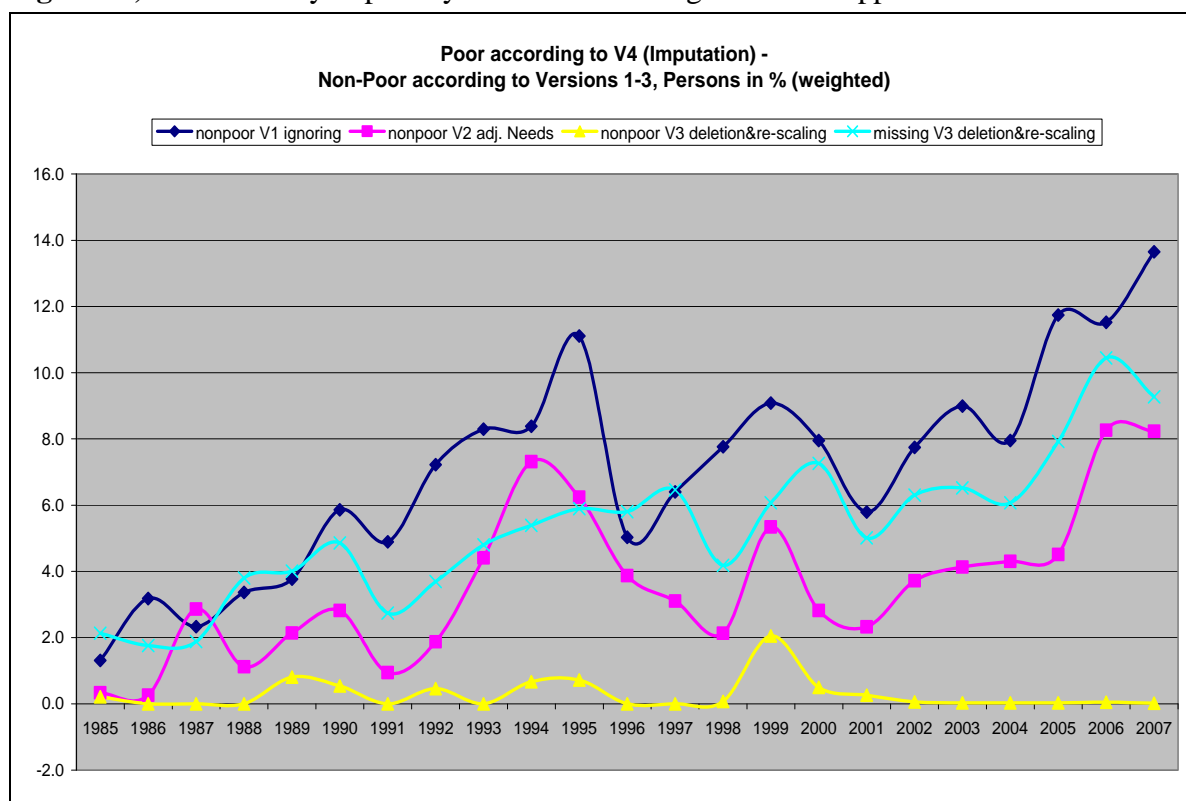


**Figure 7:** The impact of PUNR treatment on relative income poverty



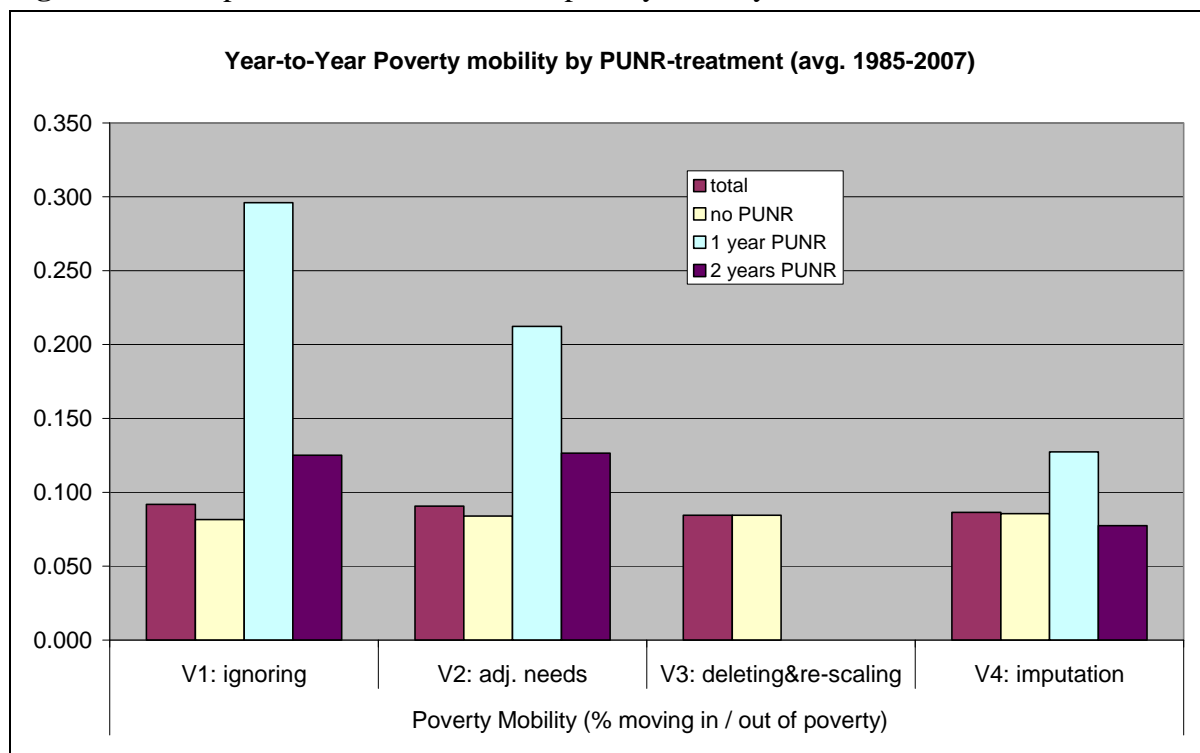
Source: SOEP. Authors' calculations.

**Figure 8a,b:** Consistency of poverty measurement using alternative approaches to handle PUNR



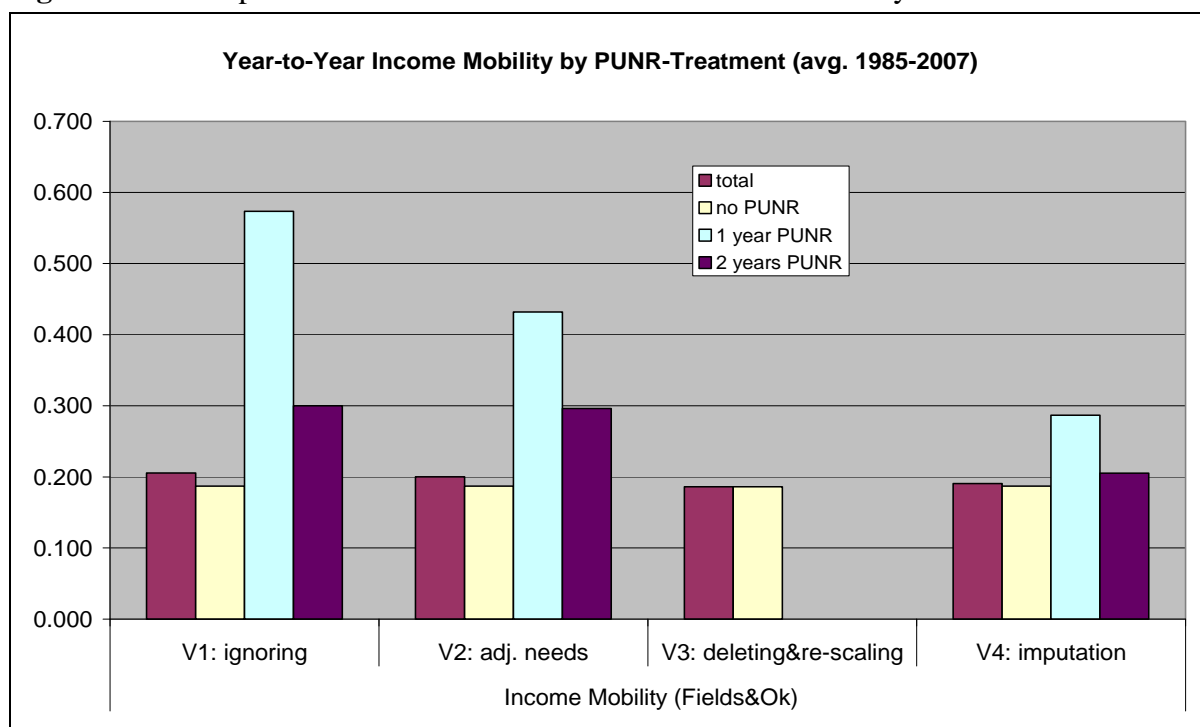
Source: SOEP. Authors' calculations.

**Figure 9:** The impact of PUNR treatment on poverty mobility



Source: SOEP. Authors' calculations.

**Figure 10:** The impact of PUNR treatment on two-wave income mobility



Source: SOEP. Authors' calculations.



**Table 1: Probability of PUNR – Results from a pooled probit regression**

		Coeff.
sex	female	-0.443 ***
age groups (ref. 56-65 yrs of age)	17-24	-0.003
	25-40	-0.053 *
	41-55	-0.210 ***
	66 and over	-0.304 ***
relation to household head (ref: household head)	partner	0.977 ***
	child	0.984 ***
	other	1.700 ***
migration background (ref: autochthonous)	native-born foreigners	0.204 ***
	foreign-born Germans	-0.264 ***
	foreign-born foreigners	-0.347 ***
East Germany (ref: West Germany)	East Germany	-0.245 ***
change in household composition (ref: no change)	change	0.076 ***
	first household interview	0.528 ***
no. of adults in household (ref: 2)	3 adults	0.052 *
	4+ adults	0.125 ***
no. of children in household (ref: none)	1 child	-0.078 ***
	2 children	-0.163 ***
	3+ children	-0.170 ***
home owner (ref: tenant)	home owner	-0.051 **
highest education of hh head/partner (ref: intermediate)	lower secondary	0.039
	higher secondary	-0.102 ***
	tertiary	-0.124 ***
item non-response monthly household income (ref: no)	yes	0.396 ***
Control dummies for survey year		yes
Constant		-1.779
Pseudo-R2=		0.1568
Observations / Individuals		325414 / 45038
Robust standard errors. * significant at 10%; ** significant at 5%; *** significant at 1%		
Source: SOEP 1985-2007. Authors' calculations.		

## Appendix A: Excerpt from the ECHP documentation on the PUNR adjustment procedure

### 6. WITHIN HOUSEHOLD NON-RESPONSE INFLATION FACTOR (HI010)

In order to have comparable figures at aggregated household level, it is necessary to make up for unit non-response if no questionnaire was answered by some persons in the household.

Remark: HI010 can only be calculated after the weighting procedure.

To do so, all persons in a country are grouped into 110 groups G - 11 age classes, sex and quintiles of equivalised net monthly household income (HI200 / HD005). For each group the weighted average (PINMEAN) of the provisional total personal income (PI100pr) is calculated.

$$PI100pr = \frac{PI111 + PI112 + PI12 + PI121 + PI123 + PI131 + PI1321 + PI1322 + PI133 + PI134 + PI135 + PI136 + (HI22 + HI137 + HI138) / HD002A}{HD002A}$$

$$PINMEAN_G = \frac{\sum_{all\ persons\ in\ group\ G} PI100pr * PG002}{\sum_{all\ persons\ in\ group\ G} PG002}$$

For households in which not all the eligible persons answered, HI010 is calculated as

$$HI010 = \frac{\sum_{all\ persons\ with\ PENRES=1,1,1\ and\ HI0} PINMEAN_G}{\sum_{all\ persons\ with\ PENRES=1\ and\ HI0} PINMEAN_G}$$

For those household where the factor is 5 or more, it will be set to missing in order to avoid that the estimated total household income is based on less than 20% of real data:

If HI010 > 5 then HI010 = -9.

Source: Eurostat (2000)

## Endnotes

- 
- <sup>i</sup> This approach is in line with the unitary model proposed by Becker (1991), which assumes a set of coherent preferences across all household members. However, it should be noted that there is a substantial body of literature questioning the general validity of this assumption (see, e.g., Vogler & Pahl 1994, Bonke & Uldall-Poulsen 2007).
- <sup>ii</sup> For a comprehensive discussion of non-response behavior in household surveys in general as well as in panel surveys in particular, see the various contributions in the readers edited by Groves & Cooper 1998, Groves et al. 2002, De Leeuw et al 2008 and Lynn 2009.
- <sup>iii</sup> It should be noted that all of the following analyses refer to the population in private households only, i.e., we exclude individuals living in institutions such as nursing homes.
- <sup>iv</sup> Without rescaling of weights, this approach is also known as “listwise deletion”. It is well-known from the literature that such simple deletion of incomplete cases without any subsequent correction most likely biases the results and increases the inefficiency of the statistics by reducing the sample size (see, e.g., Barceló 2008 for the Spanish Survey of Household Finances, Frick & Grabka 2005 for the SOEP, Frick & Grabka 2010 for the UK BHPS, the SOEP and the Australian HILDA Survey, as well as Starick & Watson 2007, also using HILDA data).
- <sup>v</sup> See Nicoletti and Peracchi (2006) for a discussion of methodological problems arising from the flat-rate correction in the ECHP.
- <sup>vi</sup> In line with the Canberra Group (2001) recommendations, we also add a measure of net imputed rent, which captures the implicit income advantage of owner-occupied housing as well as any non-monetary income advantage of subsidized renters (see Frick & Grabka 2003).
- <sup>vii</sup> According to German tax law, married couples file their taxes jointly, while all other individuals are single filers.
- <sup>viii</sup> This two-step procedure is necessary in order to avoid imputing too-low incomes for too many individuals (“regression to the mean”), in particular in cases of rarely received income components like maternity leave benefits or student aid.
- <sup>ix</sup> The quality of correction by means of imputation and weighting may crucially depend on the available data collected on the missing persons by means of proxy information (see, e.g., BHPS). A lack of such proxy data is not always the fault of the survey designers but may simply result from ethical and legal restrictions on the collection of proxy data on individuals who are unwilling or unable to participate in a survey. In fact, in the German context and thus in the SOEP, collecting proxy information on the income situation of non-participating household members is by and large not feasible. But even if the collection of proxy information were allowed, one would expect a relatively strong misspecification bias arising from the restricted knowledge base of the reporting individual with respect to the missing income information of the respective PUNR.
- <sup>x</sup> **Demographic variables** (*D=dummy variable*): age groups, family and household type, number of children, at least one household member is in need of care (D), relationship to head of household, region, community size, SOEP-subsample identifier, change in household composition (D); **Social structure**: home owner (D), (highest) education of head of household or spouse/partner, education of children, migration background

---

(D); **Income variables:** net household income at month of interview, public transfers (housing subsidies, social assistance, etc.), aggregated observed individual incomes as a share of net household income at month of interview; **Filter(s)** considering receipt of other income components (D).

- <sup>xi</sup> In an evaluation of various imputation methods, Starick (2005) argues in favor of the Little and Su methods rather than the standard (single) imputation techniques. Furthermore, the Little and Su methods perform better in maintaining cross-wave relationships and income mobility. This finding is also confirmed by Frick and Grabka (2005) for the SOEP.
- <sup>xii</sup> In more detail, the income information is weighted by  $2^{(3-dt)}$ , with  $dt$  denoting the distance in years, ranging from 1 to 3.
- <sup>xiii</sup> L&S is also the standard procedure for imputation of item non-response (INR) on income questions in the SOEP (see Frick & Grabka, 2005).
- <sup>xiv</sup> Note that in the procedure implemented here,  $r$  is based on weighted income information as described above.
- <sup>xv</sup> All empirical results presented in this paper are based on calculations using Stata (version 9.2). We gratefully acknowledge Stephen P. Jenkins and Philippe van Kerm for their Stata add-ons INEQDECO, INEQUAL7, IMOBFOK, FOKMOB, and SHORMOB.
- <sup>xvi</sup> Note that the very strong increase in the HSCV since 2003 is predominantly caused by the incorporation of SOEP's high income sub-sample G.
- <sup>xvii</sup> These findings for a two-year interval are confirmed when analyzing 5-year intervals. The lowest degree of income mobility among PUNR households can be found when using imputed data instead of applying Versions 1 or 2, while "elimination" of PUNR households from Version 3 yields the lowest degree of mobility for the entire population.
- <sup>xviii</sup> Future extensions may consider *multiple* imputation (Rubin 1987), which would also take into account the uncertainty embedded in the single two-stage imputation procedure applied here.